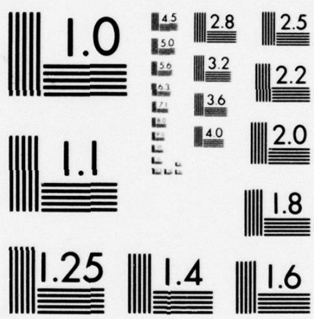
AD-A069 851 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 15/7
A COMPARISON OF TAGSEM AND RED FLAG.(U)
MAR 79 D A HIMES
UNCLASSIFIED AFIT/GST/MA/79M-3 NL

1 OF 1
AD
A069851

END
DATE
FILMED
7 -79
DDC


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

A COMPARISON OF TAGSEM
AND RED FLAG

THESIS

AFIT/GST/MA/79M-3

David A. Himes
Captain USAF

UNITED STATES AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY
Wright-Patterson Air Force Base, Ohio

A COMPARISON OF TAGSEM
AND RED FLAG

THESIS

David A. Himes
Captain USAF

AFIT/GST/MA/79M-3

(14) AFIT/GST/MA/79M-3

(6) A COMPARISON OF TAGSEM AND RED FLAG,

(9) Master's THESIS,

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(12) 43p.

by

(10) DAVID A. HIMES

~~Captain USAF~~

Graduate Strategic and Tactical Sciences

(11) Mar~~ch 1~~979

Approved for public release; distribution unlimited.

012 225

xlt

PREFACE

This research was performed to determine if an operational exercise, Red Flag, could be used to calibrate a measure of effectiveness of the Tactical Air-to-Ground System Effectiveness Model (TAGSEM). If the results of this study are valid, then the users of TAGSEM will have gained some confidence in the validity of their model.

I wish to thank Major Kenneth Melendez for his considerable help in this effort. Also, I would like to thank Mr. Paul Belcher, Mr. Timothy Ringler, and Mr. John Kordik of the Aeronautical Systems Division Deputy for Development Planning for their assistance in the problem formulation. Mr. Jerry Jensen and Mr. Robert Hubler of the University of Dayton Research Institute provided invaluable assistance in exercising the TAGSEM computer model. I gratefully acknowledge the assistance of my typist, Ms. Diane Hartzell, without whose patience this study would have been impossible.

DAVID A. HIMES

RE: AFIT/GST/MA/79M-3
Classified references, distribution unlimited-
Delete classified references per Ms. Donna Schuh, AFIT/EN

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

This research was conducted to compare a measure of effectiveness common to two models. One model, the Tactical Air-to-Ground Systems Effectiveness Model, Version II (TAGSEM II), is a computer simulation program. The other model is a field exercise, Red Flag 78-7.

Selected missions from the Red Flag exercise were simulated by the TAGSEM program to compare aircraft attrition due to ground-based defensive systems. Hypotheses of equal means were not rejected and a hypothesis of no relationship was rejected.

I INTRODUCTION AND BACKGROUND

Modeling of military activities is probably as old as recorded history if one uses Shannon's definition of a model.

> A model is a representation of an object, system, or idea in some form other than that of the entity itself (Ref: 4:4).

Various types of simulation models are currently in use for a wide variety of purposes. Two examples, each of a different type, were selected for comparison since they attempt to model similar "real-world" situations. Before proceeding further, a definition of simulation as used in this paper is necessary.

> Simulation is the process of designing a model of a real system and conducting experiments with this model for the purpose either of understanding the behavior of the system or of evaluating various strategies (within the limits imposed by a criterion or a set of criteria) for the operation of the system (Ref: 4:2).

There is presently a great deal of emphasis on simulation of all sorts, and as the processes of interest grow more complex, the managers of those processes are turning to simulation more frequently. The cost (in fiscal or other terms) of operating real systems extensively for the purpose of education or analysis is prohibitive for many activities, especially within the military. Two of the most common simulation model types are the field exercise and the computer simulation model. These models lie at almost extreme ends of the modeling spectrum with regard to level of abstraction. In most cases the more abstract tend to cost less (again, in many ways), so the field exercise and computer model both offer advantages and disadvantages.

Many computer simulation models are structured around events. Each event is represented by an instantaneous event in simulated time. When the model determines that a programmed set of conditions are met, the event occurs (Ref: 4:120). There are two basic methods of determining when a condition is met. One is Monte-Carlo sampling in which the probability of an occurrence is described by a known or estimated probability function. A random number is then generated and compared to the probability function to determine which event occurred or whether one occurred at all. Another method is to compute the expected value of the probability function and use that throughout the model. The latter has the advantage of simplicity and uses much less computer time, but it has a disadvantage in that real-world events caused by extreme values of some variable are lost.

Field exercises and maneuvers are the least abstract of all model types and tend to be primarily used for non-analytical purposes such as training or engineering development. Shannon offers a brief description of an analytical field test for the interested reader (Ref: 4:231-233). The detail and realism in a given field exercise is only limited by its designers' imaginations and their budget. Since the budget almost always runs out prior to the imagination, field exercises tend to be relatively infrequent and limited in scope.

One of the models considered is a computer simulation program called the Tactical Air-to-Ground Systems Effectiveness Model, Version II (TAGSEM II). More details of this model will be offered later, but it is basically an expected value event-oriented model (Ref: 6:89). The other model is an operational training exercise, Red Flag 78-7 (Ref: 3)

which may be considered an iconic model. This model will also be discussed later, briefly it is an operational exercise using a mixed force of actual aircraft flying against a variety of simulated and real ground targets (Ref: 1:43).

Purpose of Study

As yet, no attempt had been made to calibrate TAGSEM by any means. Due to its relevance and availability, Red Flag appeared to be the most likely candidate for a source of operational data. However, the reader should recall that as an operational training exercise, Red Flag itself is a model.

This study will compare a primary measure of effectiveness (MOE) from both TAGSEM and Red Flag. Statistical tests will be applied to determine whether or not the two models' MOE's may be considered to come from the same population and the degree of correlation between the models for those MOE's. The MOE selected was aircraft attrition due to ground-based defensive systems. Others, such as targets destroyed and weapons delivered, were considered, but the data currently collected during Red Flag exercises precluded their use.

Limitations

In general, most limitations in this study are a result of the different primary objectives of the two models. TAGSEM is designed to allow an operator to quickly examine the relative merits of tactical air-to-ground attack systems (Ref: 6:88), whereas Red Flag is primarily a training exercise for combat-ready tactical units (Ref: 1:41). Discussion of specific limitations will follow during detailed examination of the models and their integration.

The balance of this paper will describe the background for this study, detailed descriptions of the two models, integration of the Red Flag data base into the computer model, a statistical comparison of results, and a discussion of some factors that affect this comparison.

Background

In an effort to provide an analytical tool for evaluating various combinations of attacking aircraft, weaponry, and enemy threats, the Deputy for Development Planning, Aeronautical Systems Division (ASD/XRO), in cooperation with the Deputy for Tactical Planning (ASD/XRP), developed TAGSEM II. This model is used by agencies involved in analyses pertaining to force mixes, offensive and defensive weapon systems, and strike tactics, to name a few. Generally, TAGSEM is designed to accomplish one of the basic functions of a model, that is, to evaluate the effects of various strategies (Ref: 4:2). To date, no quantitative measure of the validity of TAGSEM has been made. Since TAGSEM's purpose is to provide a measure of relative, not absolute, merit (Ref: 6:88), an attempt to validate the model to the real world was thought to be unnecessary. It was, therefore, decided to find another "yardstick" by which an attempt to calibrate part of the model might be possible.

The Red Flag exercises conducted by the 4440th Tactical Training Group (TTG) are intended to simulate as nearly as possible a wartime environment. Analyses of the air war in Southeast Asia (SEA) and other conflicts showed the incidence of combat losses dropped considerably after the pilot had completed about ten sorties. The primary mission of Red Flag is to provide U.S. tactical aircrews that combat experience in a

peacetime environment (Ref: 1:41). In a sense then, Red Flag fulfills the other purpose of simulation mentioned previously, that is, it increases the understanding of the system in question (Ref: 4:2).

Red Flag exercises have some important features in common with TAGSEM that other operational exercises (or even well-documented previous conflicts) do not. Perhaps the most apparent is the absence of ground troops. Although this makes the conduct of a combat air operations somewhat artificial, one must recall the purposes of the two models. Neither is intended to trace or predict the movement of the Forward Edge of the Battle Area (FEBA) as real estate is gained and lost by the competing ground armies. Although this factor would have immense impact on actual air operations, both models are much more limited in scope. Another feature is that both models launch strike sorties into a target array predetermined by the user. Although Red Flag does not actually shoot down aircraft or destroy ground targets, parameters relating to the success or failure of missions are measured.

For these reasons, it was thought that Red Flag bore sufficient similarity to TAGSEM to allow TAGSEM to be used as a computer simulation of the Red Flag operational exercise. As will be pointed out later, there are also many dissimilarities, but these do not preclude preparing a set of input data for the computer model.

The chapter that follows will describe in some detail both models. Although the reader will note many dissimilarities between them, he should recall the MOE of interest. Many aspects of both models are not considered in this study.

II The Models

This chapter is a summary of information contained in the TAGSEM User's Guide and the Red Flag final report (Ref: 3 and 6). The principles of operation for both models are outlined and the reader is referred to the above references for further information.

TAGSEM II

This model is an expected value computer program that evaluates the relative effectiveness of tactical air-to-ground weapon systems. It is written in FORTRAN IV and is presently run on the Control Data Corporation 6600 computer with the NOS/BE operating system (Ref:6 :4). The forces involved may range from a single aircraft attacking a single target to a mixed force of attackers including support aircraft striking a large complex array of targets. The key to this model's simplicity is that the probabilities associated with each event are determined off-line by other models, empirical data, or subjective judgment. For each aircraft in the attacking force the model uses probabilities associated with factors such as scenario, aircraft performance, mission profile, payload, survivability, target acquisition capabilities, weapon lethalities, navigation accuracies, and sortie rate in order to evaluate the overall effectiveness of a particular type of aircraft. The effect of various types of support aircraft may also be studied. TAGSEM makes no attempt to determine cost.

TAGSEM computes primary measures of effectiveness (MOE) as a function of time. Some of these MOE's (all expressed as expected values) are: each type of aircraft lost and damaged, each type of target destroyed by each type of aircraft, weapons delivered on each type of target by

each type of aircraft, sorties flown, and enemy air defense sites destroyed by type.

The computer model may simulate a mixed force of attack and support aircraft with the payload of each aircraft type specified by the user. The user also specifies the mission profile including the target array to be attacked. This array can consist of numerous target types such as tanks, trucks, and armored personnel carriers (APC). Enemy surface-to-air missile (SAM) and anti-aircraft artillery (AAA) units may be among these targets. The user may position additional SAM and AAA threats on the ingress and egress routes.

At this point it may be useful to illustrate what a typical "battlefield" could look like. Figure 1 shows a possible scenario with ten zones, although the user may elect to use any smaller number.

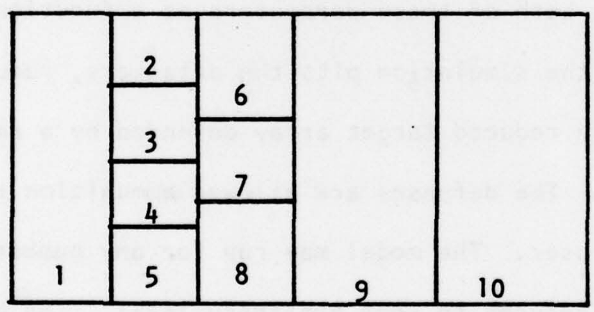


Figure 1

There is no requirement for symmetry. The size of each zone in square kilometers is input to the model along with the number and type of all targets and defensive weapon systems in each zone. The distribution within each zone is assumed uniform. It is also possible to specify the ability of each SAM or AAA weapon to fire into other zones. This

recognizes the fact that the range of some defensive systems exceeds the size of any one zone (or even the entire area of interest). The user provides the percentage of each defensive system, by type and zone, that may fire into each other zone. For example, 30% of a particular type of defensive system in zone 5 (see Fig. 1) may fire into zone 8, 10% each into zones 7 and 9, 30% each into zones 4 and 6, 70% into zone 3, and none into the other zones. The total percentage may exceed 100%, although no more than 100% may be active at any one time.

After surviving aircraft return to base, those with battle damage enter into a repair cycle. There are two types of damage, repairable and nonrepairable. In the case of repairable damage, the user specifies the average down time by aircraft type and defensive system. Nonrepairable damage is that which takes longer to repair than the length of the "conflict". The user specifies both of these parameters as a function of aircraft lost. The next cycle of the simulation pits the attackers, reduced in number, against a similarly reduced target array defended by a reduced number of defensive systems. The defenses are allowed ammunition re-supply at a rate determined by the user. The model may run for any number of cycles or until one side is reduced to some arbitrary level. The MOE's may be printed out by cycle or in summary.

Some limitations that affect the computer model in general and this study in particular should be addressed. The most obvious area ignored is command and control. There is no direct simulation of a command control network, but input data may be altered to reflect the effect of such a system. Also, there is no simulation of a ground battle or air-to-air conflict. These effects may be included in off-line determination of input data, but they cannot be varied during the actual simulation run.

One other area that inhibits TAGSEM's flexibility is its scenario limitations. Although a great many types of strike and support aircraft may be employed against a fairly complex target array, the same strike force must continue to attack the same target array in the same fashion. The computer model will attrit both forces during the simulation, however, the user may not make external changes to the scenario such as reinforcement or targeting changes (Ref: 6:89-93).

Red Flag

As was previously pointed out, the primary purpose of Red Flag exercises is to give operational aircrews practical experience under conditions as near to combat as possible. These exercises are conducted on the range complex near Nellis AFB, Nevada. Each exercise centers around an operational tactical unit, usually a squadron, that is deployed to Nellis AFB for the period of operations, normally about 30 days. A secondary benefit is the experience gained by the maintenance, logistic, and intelligence personnel deployed in support of the flying operations. Supporting aircraft are flown out of Nellis and other locations by other tactical units, Major Air Commands (MAJCOM), as well as U.S. Navy and allied air units.

Most of the ground targets used in the exercises are actual or mock-up tanks, trucks, aircraft, etc., that are placed on the range under realistic deployment situations. For instance, one convoy uses real trucks spread along 17 miles of road with .8 mile spacing. There are no bombing circles or cloth panels such as those used in gunnery training.

Included in the range are simulated SAM and AAA radars and communications jamming equipment. Although no missiles are actually launched or

guns fired against attacking aircraft, the radar tracking information is stored and may be viewed on a television screen showing the aircraft flight path and the successes or failure of radar tracking.

A great deal of effort is devoted to air-to-air training in both offensive and defensive situations. Again, no ordnance is fired, but videotapes of radar tracking information are used in analyses of engagements.

Two things that Red Flag does not do is exercise the tactical air control system or the movement of ground armies. Although there is a complex mix of aircraft flying a wide range of missions during Red Flag exercises, each day of the "war" starts with no effects or constraints from previous activities. Gaining and losing real estate, effectiveness of command and control, and other similar factors would obviously bear heavily on the outcome of an actual conflict, however, the reader should bear in mind that the primary purpose of Red Flag is to train aircrews and not to simulate all aspects of a conventional conflict (Ref: 1:41-43).

How these two models were combined is the next area covered. Many of the assumptions made will be noted as well as some of the characteristics of the computer model that affected the type of data extracted from the Red Flag report.

III INTERFACE OF THE MODELS

A discussion of the TAGSEM-Red Flag interface follows. The reason for selecting the MOE of interest as well as the method of computer simulation will be covered. The remainder of the chapter will describe the actual interface of the models.

As has been previously stated, the only MOE compared between the two models was aircraft attrition due to ground-based defensive systems. TAGSEM produces a very wide variety of MOE's, as do most models used for analysis. Almost any operational parameter for air-to-ground attack is available in one form or another. However, Red Flag has primarily a training objective, and its MOE's of interest are different. An important point to bear in mind is that real people and aircraft participate in a Red Flag exercise, which makes it very expensive. Exotic range instrumentation and data collection techniques that are superfluous to the primary mission are not available. Also, a great deal of effort is devoted to air-to-air combat training and tactics, so considerable resources are spent on data collection in this area.

In the conceptual phase of this study, there was some interest in evaluating MOE's such as targets destroyed or weapons on target. Although the simulated type of ordnance and the target for each aircraft in a Red Flag exercise is available, the only information concerning weapon-target interaction is miss distance. For point targets the weapon effects would be highly probabilistic and even more so for area targets. Perhaps weapon effects on point targets could be compared, but the distributions of

vulnerable points in area targets is unknown. For that matter, in some cases such as an airfield complex, the definition of what is a vulnerable point would be open to debate.

Using similar arguments, the concept of "on target" is of little value unless one has considerable knowledge of weapon-target interaction again. "On target" when a tank is the target is very different than if it were a factory. Since an independent study in weaponeering is beyond the scope of this paper, the one commonly tracked MOE, aircraft attrition, was selected for investigation. In a Red Flag exercise, the concept of "validated attack" is used. When a simulated defensive weapon meets its preset launch parameters, a validated attack is logged against the target aircraft. No weapons effects are considered.

This analysis was performed by making numerous small-scale TAGSEM runs representing individual Red Flag missions. The other possible approach would have been to model the entire exercise on one or two large-scale computer runs. Several factors mitigated against this approach. As indicated in Chapter II, once the data for a particular scenario is input, the computer model runs that same scenario until completion. Since there is no replacement of attacking aircraft, defensive systems, or targets, subsequent cycles will play fewer and fewer elements. In an actual Red Flag exercise each mission is planned and flown operationally independent of all others. Also, the actual number of various aircraft types varies sufficiently from mission to mission so as to exceed the capacity of TAGSEM to track aircraft types. For these reasons TAGSEM was set up to model many small independent missions.

The remainder of this chapter will describe each variable in the TAGSEM model as it is used to reflect Red Flag data. Those variables that change from one run to another for the different Red Flag missions that are examined will be pointed out. For a detailed description of input data format, the reader should refer to the User's Guide (Ref: 6).

The first information block concerns the force levels involved such as the number of types of defenses, zones, targets, and aircraft. The number of target arrays is also included. All of these inputs are changed for each computer run. The number of cycles per repetition and number of repetitions is one in all cases. Since single missions are being considered, the model allows for less than perfect bomb damage assessment (BDA) for attacking aircraft, however, since each Red Flag mission starts with a clean slate, the value used for all runs is 1.0. The computation of the probability that a defensive system will fire on an attacker is computed iteratively, and all runs use the same values for these variables that the model developers typically use (20).

Switches that control the various outputs are set to either one or zero in order to print or suppress information. Only summary tables were obtained, although a user may elect to observe each event if he so desires.

Information describing the defensive systems is now entered. For this study, five system types were modeled. Each type is identified by the following data: kind of system (SAM or AAA), target number (recall that each SAM or AAA piece may also be a target for attacking aircraft), rounds available at each site, and number of systems in each zone. The computer model assumes uniform density within each zone. Since the Red Flag threat radars are not targeted for safety reasons, no target number is provided.

Arbitrarily high ammunition stocks were used to allow the defender to "fire" in an unconstrained manner.

Similar data is then provided for each aircraft type, however, "identification of escort" is a slightly confusing concept. Some aircraft may be limited to working exclusively in concert with another type (consider a hunter-killer team). Therefore, each aircraft must have its escort identified for the computer model. For the majority of the aircraft of interest they are their own escort. That is, they are allowed to operate without any other type of aircraft present. The general type of aircraft (attack, observation, etc.) and the number of sorties per cycle it may fly are entered. This study only considered attack aircraft. Next a series of inputs defining force size by aircraft type and how that force may be subdivided is defined. Each Red Flag mission has its own characteristics in this area, but as a matter of practice many of these variables do not change from mission to mission.

The interaction of aircraft damaged versus defenses are provided next. In a simulation run consisting of only one attack, the difference between repairable and nonrepairable damage is of little value, but for computational purposes these inputs were provided. Both types of damage are computed as some constant times the number of aircraft lost. The average repair time for repairable damage is also specified. For each aircraft type and defense type these simulation runs use the same coefficients as typically used by the model developers. The model has the capacity to play reduced visibility, however, this aspect was ignored and unrestricted visibility factors were used.

The target arrays are next defined. Each target array is numbered and the zone in which it is located specified. The type and number of each target and the type and number of defenses within the array are entered. Since the active radar defenses at Red Flag are not located within the target area, there are zero entries here for defenses. The specific target or targets attacked in each Red Flag mission will determine the other elements.

A rather complex set of information that has little bearing on this particular application is next and is discussed here mainly to give the reader a better understanding of the computer model. The activities of each cycle, the number of which has been previously specified, are now defined. An integer identifying an event set is associated with each cycle. For this application there is only one cycle for each Red Flag mission, however, the event set that plays the aircraft through their individual sorties is still necessary. The number and type of each aircraft in each cycle is provided and now each one of these aircraft types is associated with an event or events within the cycle. There may be a number of events in a cycle with any or all of the appropriate aircraft participating in each event. The number of events and number of aircraft types in each event are determined by the user's judgment as to the activities he is trying to simulate. The activities of each aircraft type during each event are specified by a sequence of integers, each one associated with a code provided in the User's Guide. A zero entry indicates that a particular aircraft does not participate in that particular event, and an entry such as 1024 would indicate that in that event the activities specified in various library segments (to be discussed later) are performed.

The user may elect to allow redistribution of defensive weapons between cycles. If allowed, the surviving defensive systems will be repositioned uniformly throughout the battle area. The user may also allow re-supply of ammunition at a rate of his choosing. Neither of these activities were played.

Each aircraft type has a series of library segments that define various aspects of its mission profile. The mission is broken down into subsegments which position the aircraft in different zones. It is here that the area of the zones and the distance that each aircraft travels through the zone is defined. The ability of defenses not in the zone through which the aircraft is flying to engage the aircraft is stated for each defense type by zone and mission segment. Only egress through a zone of defenses was played.

Each defense type that appears in a library segment (and is thus mated with a particular aircraft as a possible target) has a series of probabilities associated with it. These are: that it is operationally ready at the time of the encounter, that the target can avoid its lethal envelope, its reliability, probability of a kill given a firing, and expected number of rounds fired per encounter. The ground ranges of the weapons are listed here. For the runs made in this study, the same values for these variables were used as commonly used by the developers of the model. The probability that an aircraft will not abort prior to engagement is in this data block, and for these runs 1.0 was always used. Since the mission final reports stated any aborts, only those aircraft that did not abort were played on the computer model.

A large section in each library segment is devoted to terminal effects. Since there are no terminal defenses at Red Flag, all probabilities associated with their effects are zero. At this point the condition of the target (always new for each Red Flag mission), the number and combination of weapons delivered (specified for each mission in the Red Flag final report), and the probabilities that the aircrew will be able to find and see the target are entered. As in the abort situation, those aircraft that actually did not find the target are not considered. Once over the target or in its immediate vicinity, probabilities associated with weapons delivery are required. Probabilities that the pilot can lock onto the target (for guided weapons) or meet visual release parameters (for unguided) on any single pass, that the target will be destroyed on the first pass, and that the target will be destroyed on any subsequent pass are provided to the model. The maximum number of passes and the maximum number of weapons releasable per pass are included. Those figures commonly used by the model developers again were used for particular weapons dropped or simulated at the Red Flag exercise. Coefficients for weighting target values to the aircrews were available, however, in this application target systems were very simple, so unity was applied throughout. Even though target attack was not played, this data is required to run the model.

After the simulation runs were complete, the aircraft attrition for the two models was compared. The following chapter describes the tests used and their results.

IV RESULTS OF THE EXPERIMENT

A test of the means of populations which assumes normally distributed data and one not requiring such an assumption are performed in this chapter. Also, a coefficient of correlation and its use in a test for a possible relationship between two samples is described. Accompanying the calculations are various tables with supporting data.

Three aircraft types, A, B, and C were pitted against five defensive system types, 1 through 5. Aircraft types A and B are current Air Force equipment for which the required data for input was available. Type C was actually a surrogate for numerous actual aircraft types for which off-line analyses have not been performed. Input data characteristics for aircraft A were used since this appeared to offer a compromise in performance among the various types involved. Aircraft C was included to allow for dilution effects of a large number of aircraft on the defenses. Aircraft not involved in egress through the defensive system were not played. Red Flag missions were selected that had similar profiles so as to allow a single scenario on the computer simulation model to be used. Table 1 shows the number of aircraft involved and their attrition.

A useful technique when comparing two alternatives is correlated sampling. This technique eliminates a source of the extraneous variance that may occur from pair to pair. This is done by calculating the variance of the difference of pairs of observations rather than the variance among the individuals.

A summary of aircraft attrition is tabulated on Table 2 by aircraft type versus defense type. The hypothesis of no difference between the

population means was tested using the paired observations technique as described by Shannon (Ref: 4:222-224). The test statistic t was calculated as

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\Sigma D^2 - (\Sigma D)^2/N}{N(N-1)}}} \qquad (1)$$

where $\overline{X}_1$ = Red Flag average aircraft lost due to each defense type

$\overline{X}_2$ = TAGSEM average aricraft lost due to each defense type

D = difference between each paired observation

N = number of paired observations (15)

This yielded a calculated value for t of 1.404. At $\alpha = .05$ the tabulated value of t for a two tailed test at 14 degrees of freedom is 2.14. The hypothesis of no difference between population means is not rejected.

Mann-Whitney Test

Use of the Student-t distribution requires an assumption of normality (Ref: 5:135) and it is difficult to determine if this is indeed the case. Therefore, as a check the nonparametric Mann-Whitney test as described by Miller and Freund (Ref: 2:275-278) was used also. Since the number of observations is sufficiently large (over 8), the statistic

$$z = \frac{\frac{n_1 n_2 + n_1 (n_1+1)}{2} - R_1}{\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}} \quad (2)$$

where $n_1 = n_2$ = number of observations (15)

R_1 = sum of TAGSEM or Red Flag ranks

was calculated. This yielded a value for z of plus or minus .56 depending on which sum of ranks was used. Since $n_1 = n_2$, either sum may be used. At α=.05 for a two tailed test z equals 1.96. Again, the null hypothesis that the means are equal is not rejected.

Correlation Coefficient

The calculation of correlation coefficients as described by Shannon (Ref: 4:86-87) was performed on both the sample as a whole and on individual samples with respect to aircraft type. The coefficient is given by

$$r = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[n\Sigma X^2 - (\Sigma X)^2][n\Sigma Y^2 - (\Sigma Y)^2]}} \quad (3)$$

where n = number of samples

X = TAGSEM losses

Y = Red Flag losses

The correlation coefficient for all aircraft types versus each defense type is .647. For aircraft A against each defense it is .553, for B . 628, and for C .898. (See Table 3).

Miller and Freund offer a test (Ref: 2:325) to determine if there is no relationship between two samples using the correlation coefficient. The statistic

$$z = \frac{\sqrt{n-3}}{2} \times \ln \frac{1+r}{1-r} \qquad (4)$$

where n = number of samples

r = correlation coefficient

was used to test the null hypothesis that there is no relationship between TAGSEM and Red Flag aircraft attrition. Using the previously determined values for r, the z statistic for all aircraft types versus each defense is 2.67, for only type A against each defense is .88, for type B is 1.04, and type C is 1.61. For a two tailed test at α=.05, the tabulated value for z is 1.96. The null hypothesis of no relationship between the two samples is not rejected for the individual aircraft types, however, it is rejected for all types considered together. That is, the alternate hypothesis that there is a relationship is accepted.

Since the purpose of Red Flag centers on training, one would expect (for the MOE of interest) fewer combat losses on late missions than on early ones. Tables 4 and 5 show the same information as Table 2 except that Table 4 is for the first 8 missions (chronologically) and Table 5 is for the last 8. Using Equation (1), the test statistic for early missions is 1.604 and for late is -2.23. At α=.05 the null hypothesis is not rejected for the early missions and is rejected for the late missions. It would seem that the number of aircraft lost late in the exercise is less than the expected number of aircraft losses from the computer model.

The next chapter will point out many of the specifics about both models that have not yet been covered. Some sources of possible error due to assumptions and simplifications will be discussed.

AIRCRAFT TYPE A

NUMBER LAUNCHED	1	2	3	4	5	DEFENSE TYPE
8	0 .02	0 .04	3 .08	0 .064	0 .09	Red Flag Losses* TAGSEM Losses
7	0 .02	0 .04	1 .08	0 .064	1 .09	
11	0 .02	0 .04	1 .081	0 .065	0 .091	
13	0 .02	0 .04	0 .081	1 .065	0 .092	
8	0 .01	0 .02	3 .041	0 .033	1 .046	
8	0 .01	0 .02	4 .041	0 .033	0 .046	
4	0 .01	0 .02	0 .04	0 .032	0 .045	
6	0 .02	0 .039	0 .079	0 .032	0 .09	
4	0 .01	0 .02	0 .04	0 .032	0 .045	
3	0 .01	0 .02	0 .04	0 .032	0 .045	
4	0 .01	0 .02	0 .04	0 .032	0 .045	
5	0 .02	0 .039	0 .079	0 .064	0 .09	
7	0 .02	0 .04	0 .08	0 .064	0 .09	
4	0 .02	0 .038	0 .078	0 .064	0 .09	
7	0 .019	0 .039	0 .078	0 .064	0 .089	
9	0 .019	0 .039	0 .079	0 .064	0 .089	
3	0 .019	0 .039	0 .079	0 .064	0 .089	

*Actually validated attacks, i.e., weapon meets its launch parameters.

TABLE 1a

AIRCRAFT TYPE B

NUMBER LAUNCHED	1	2	3	4	5	DEFENSE TYPE
7	0 .018	0 .036	0 .032	0 .026	0 .037	Red Flag Losses TAGSEM Losses
0						
4	0 .008	0 .016	0 .032	0 .026	0 .037	
4	0 .008	0 .016	0 .032	0 .026	0 .037	
0						
0						
0						
0						
0						
2	0 .01	0 .019	0 .039	0 .032	0 .045	
4	0 .008	0 .016	0 .032	0 .026	0 .037	
2	0 .01	0 .019	0 .039	0 .032	0 .045	
3	0 .01	0 .02	0 .04	0 .032	0 .045	
2	0 .01	0 .019	0 .039	0 .032	0 .045	
0						
0						
2	0 .01	0 .019	0 .039	0 .032	1 .045	

TABLE 1b

AIRCRAFT TYPE C

NUMBER LAUNCHED	1	2	3	4	5	DEFENSE TYPE
7	0	0	0	0	0	Red Flag Losses
	.01	.02	.04	.033	.046	TAGSEM Losses
6	0	0	0	0	0	
	.01	.02	.04	.033	.046	
12	0	0	1	1	1	
	.01	.02	.041	.033	.046	
15	0	0	0	0	3	
	.01	.02	.041	.033	.046	
0						
2	0	0	1	0	0	
	.01	.019	.039	.032	.045	
8	0	0	0	0	0	
	.01	.02	.041	.033	.046	
5	0	0	0	0	0	
	.01	.02	.04	.032	.046	
6	0	0	0	0	0	
	.01	.02	.04	.033	.046	
4	0	0	0	0	0	
	.01	.02	.04	.032	.045	
8	0	0	0	0	0	
	.01	.02	.041	.033	.046	
4	0	0	0	0	0	
	.01	.02	.04	.032	.045	
16	0	0	0	0	0	
	.01	.02	.041	.033	.046	
0						
0						
0						
2	0	0	0	0	0	
	.01	.019	.039	.032	.045	

TABLE 1c

DEFENSE TYPE	AIRCRAFT TYPE A	B	C	
1	0	0	0	Red Flag Losses
	.277	.092	.13	TAGSEM Losses
	-.277	-.092	-.13	D
2	0	0	0	
	.552	.18	.258	
	-.552	-.18	-.258	
3	12	0	2	
	1.114	.324	.523	
	10.886	-.324	1.477	
4	1	0	1	
	.867	.264	.424	
	.133	-.264	.576	
5	2	1	4	
	1.261	.373	.594	
	.739	.627	3.406	

TABLE 2

AIRCRAFT VS DEFENSE	TAGSEM X	RED FLAG Y	XY	X^2	Y^2
A/1	.277	0	0	.0767	0
A/2	.552	0	0	.3047	0
A/3	1.114	12	13.368	1.2410	144
A/4	.867	1	.867	.7517	1
A/5	1.261	2	2.522	1.5901	4
B/1	.092	0	0	.0085	0
B/2	.18	0	0	.0324	0
B/3	.324	0	0	.1050	0
B/4	.264	0	0	.0697	0
B/5	.373	1	.373	.1391	1
C/1	.13	0	0	.0169	0
C/2	.258	0	0	.0666	0
C/3	.523	2	1.046	.2735	4
C/4	.424	1	.424	.1798	1
C/5	.594	4	2.376	.3528	16
	7.233	23	20.976	5.2085	171

r = .647

TABLE 3

AIRCRAFT TYPE

DEFENSE TYPE	A	B	C	
1	0	0	0	Red Flag Losses
	.129	.034	.06	TAGSEM Losses
	-.129	-.034	-.06	D
2	0	0	0	
	.259	.068	.119	
	-.259	-.068	-.119	
3	12	0	2	
	.523	.096	.242	
	11.477	-.096	1.758	
4	1	0	1	
	.420	.078	.197	
	.58	-.078	.803	
5	2	0	4	
	.589	.111	.275	
	1.411	-.111	3.725	

TABLE 4

AIRCRAFT TYPE

DEFENSE TYPE	A	B	C	
1	0	0	0	Red Flag Losses
	.128	.058	.06	TAGSEM Losses
	-.128	.058	-.06	D
2	0	0	0	
	.255	.112	.119	
	-.255	-.112	-.119	
3	0	0	0	
	.514	.228	.241	
	-.514	-.228	-.241	
4	0	0	0	
	.416	.186	.195	
	-.416	-.186	-.195	
5	0	1	0	
	.583	.262	.273	
	-.583	.738	-.273	

TABLE 5

V EXPERIMENT ANALYSIS

This chapter will be classified since it deals with actual aircraft and defensive system information. Basically, how my assumptions affected the outcome of the experiment will be discussed. This chapter is available at the Deputy for Development Planning, Aeronautical Systems Division (ASD/XRO) located at Building 47, Wright-Patterson Air Force Base, Ohio 45433.

VI CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Use of an operational exercise (Red Flag) to calibrate at least one measure of effectiveness in the TAGSEM II computer model is appropriate. However, the reader should recall the many assumptions and simplifications that were used. The users of TAGSEM may gain some confidence that both TAGSEM and Red Flag to some extent are capturing the same dynamics. This statement should not be viewed as absolute for several reasons.

First, even if TAGSEM and Red Flag are modeling the same dynamics, it is possible that neither model actually captures reality. Also, the simplifying assumptions in this report may be sufficient to jeopardize the statistical inferences that were drawn. Thirdly, the statistical tests used failed to reject the hypotheses of equal means in one case or of no relationship in another which opens the door for an error of the second kind or a β error (Ref: 5:60). The data itself with many zero or near zero points should be approached with some caution. Although sensitive to sample size, the results of Equation (4) when applied to individual aircraft types should caution the reader. Apparently the computer model does not capture the attrition of individual aircraft types as effectively as it does the entire force. Also, the possible presence of a learning curve in a sample of this size may obscure the results. There is a chance that steady-state was never realized in the Red Flag exercise that only lasted about a month. The tests used to not prove beyond a shadow of a doubt that TAGSEM and Red Flag model the real-world accurately; they merely give some confidence that the two models have one effect in common.

Recommendations

As the scope of Red Flag operations widens and more comprehensive data is collected, it may prove feasible at some future date to perform a similar analysis covering more TAGSEM MOE's, especially those dealing with weapon effects. This effort was restricted to the type of Red Flag information available which was mostly in the area of aircraft attrition.

The general concept of using one model to calibrate another may be expanded in another study. Any other model should be as different in structure from TAGSEM as possible, whether or not it is a computer simulation. An interesting exercise would be to use a general purpose theatre level simulation model in a role similar to the one played by Red Flag in this study.

BIBLIOGRAPHY

1. Arnold, Terry. "Red Flag TAC's Realistic Approach to Readiness", Air Force Magazine, 60: 40-44, January, 1977.

2. Miller, Irwin and John E. Freund. Probability and Statistics for Engineers. Englewood Cliffs: Prentice-Hall, Inc., 1977.

3

4. Shannon, Robert E. Systems Simulation: the Art and Science. Englewood Cliffs: Prentice-Hall, Inc., 1975.

5.

6. -----. User's Guide for the Tactical Air-to-Ground Systems Effectiveness Model, Version II (TAGSEM II). Wright-Patterson AFB, Ohio: Aeronautical Systems Division, 1978 (Draft copy).

7. Walker, Helen M. and Joseph Lev. Statistical Inference. New York: Henry Holt and Company, 1953.

VITA

David A. Himes was born in Chickasha, Oklahoma on September 14, 1947. He graduated from high school in Bangor, Maine in 1965 and attended the United States Military Academy from which he graduated in 1969 with a Bachelor's degree in Military Engineering and a commission in the United States Air Force.

After completing pilot training at Reese Air Force Base, Texas, he operationally flew the KC-135 at Dyess Air Force Base, Texas and Southeast Asia, the OV-10 in Southeast Asia, the T-37 at Williams Air Force Base, Arizona and Randolph Air Force Base, Texas, and the T-38 at Randolph Air Force Base, Texas. He entered the Air Force Institute of Technology in August 1977.

He is married to the former Tonya L. Coombs of Mesa, Arizona. They have two sons, Brian and Nathan.

Permanent Address: 84 Lake Lorraine Circle
Shalimar, Florida 32579

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/GST/MA/79M-3	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A COMPARISON OF TAGSEM AND RED FLAG		5. TYPE OF REPORT & PERIOD COVERED M.S. Thesis
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David A. Himes		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT/EN) Wright-Patterson AFB, Oh 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Institute of Technology (AFIT/EN) Wright-Patterson AFB, Oh 45433		12. REPORT DATE March, 1979
		13. NUMBER OF PAGES 39
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

25 MAY 1979

APPROVED FOR PUBLIC RELEASE AFR 190-17.

JOSEPH P. HIPPS, Major, USAF

Director of Information

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Red Flag
TAGSEM
Model Comparison

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This research was conducted to compare a measure of effectiveness common to two models. One model, the Tactical Air-to-Ground Systems Effectiveness Model, Version II, (TAGSEM II), is a computer simulation program. The other model is a field exercise, Red Flag 78-7.

Selected missions from the Red Flag exercise were simulated by the TAGSEM program to compare aircraft attrition due to ground-based defensive systems. Hypotheses of equal means were not rejected and a hypothesis of no relationship was rejected.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

INSTRUCTIONS FOR PREPARATION OF REPORT DOCUMENTATION PAGE

RESPONSIBILITY. The controlling DoD office will be responsible for completion of the Report Documentation Page, DD Form 1473, in all technical reports prepared by or for DoD organizations.

CLASSIFICATION. Since this Report Documentation Page, DD Form 1473, is used in preparing announcements, bibliographies, and data banks, it should be unclassified if possible. If a classification is required, identify the classified items on the page by the appropriate symbol.

COMPLETION GUIDE

General. Make Blocks 1, 4, 5, 6, 7, 11, 13, 15, and 16 agree with the corresponding information on the report cover. Leave Blocks 2 and 3 blank.

Block 1. Report Number. Enter the unique alphanumeric report number shown on the cover.

Block 2. Government Accession No. Leave Blank. This space is for use by the Defense Documentation Center.

Block 3. Recipient's Catalog Number. Leave blank. This space is for the use of the report recipient to assist in future retrieval of the document.

Block 4. Title and Subtitle. Enter the title in all capital letters exactly as it appears on the publication. Titles should be unclassified whenever possible. Write out the English equivalent for Greek letters and mathematical symbols in the title *(see "Abstracting Scientific and Technical Reports of Defense-sponsored RDT/E,"AD-667 000)*. If the report has a subtitle, this subtitle should follow the main title, be separated by a comma or semicolon if appropriate, and be initially capitalized. If a publication has a title in a foreign language, translate the title into English and follow the English translation with the title in the original language. Make every effort to simplify the title before publication.

Block 5. Type of Report and Period Covered. Indicate here whether report is interim, final, etc., and, if applicable, inclusive dates of period covered, such as the life of a contract covered in a final contractor report.

Block 6. Performing Organization Report Number. Only numbers other than the official report number shown in Block 1, such as series numbers for in-house reports or a contractor/grantee number assigned by him, will be placed in this space. If no such numbers are used, leave this space blank.

Block 7. Author(s). Include corresponding information from the report cover. Give the name(s) of the author(s) in conventional order *(for example, John R. Doe or, if author prefers, J. Robert Doe)*. In addition, list the affiliation of an author if it differs from that of the performing organization.

Block 8. *Contract or Grant Number(s)*. For a contractor or grantee report, enter the complete contract or grant number(s) under which the work reported was accomplished. Leave blank in in-house reports.

Block 9. Performing Organization Name and Address. For in-house reports enter the name and address, including office symbol, of the performing activity. For contractor or grantee reports enter the name and address of the contractor or grantee who prepared the report and identify the appropriate corporate division, school, laboratory, etc., of the author. List city, state, and ZIP Code.

Block 10. Program Element, Project, Task Area, and Work Unit Numbers. Enter here the number code from the applicable Department of Defense form, such as the DD Form 1498, "Research and Technology Work Unit Summary" or the DD Form 1634. "Research and Development Planning Summary," which identifies the program element, project, task area, and work unit or equivalent under which the work was authorized.

Block 11. Controlling Office Name and Address. Enter the full, official name and address, including office symbol, of the controlling office. *(Equates to funding/sponsoring agency. For definition see DoD Directive 5200.20, "Distribution Statements on Technical Documents.")*

Block 12. Report Date. Enter here the day, month, and year or month and year as shown on the cover.

Block 13. Number of Pages. Enter the total number of pages.

Block 14. Monitoring Agency Name and Address *(if different from Controlling Office)*. For use when the controlling or funding office does not directly administer a project, contract, or grant, but delegates the administrative responsibility to another organization.

Blocks 15 & 15a. Security Classification of the Report: Declassification/Downgrading Schedule of the Report. Enter in 15 the highest classification of the report. If appropriate, enter in 15a the declassification/downgrading schedule of the report, using the abbreviations for declassification/downgrading schedules listed in paragraph 4-207 of DoD 5200.1-R.

Block 16. Distribution Statement of the Report. Insert here the applicable distribution statement of the report from DoD Directive 5200.20, "Distribution Statements on Technical Documents."

Block 17. Distribution Statement *(of the abstract entered in Block 20, if different from the distribution statement of the report)*. Insert here the applicable distribution statement of the abstract from DoD Directive 5200.20, "Distribution Statements on Technical Documents."

Block 18. Supplementary Notes. Enter information not included elsewhere but useful, such as: Prepared in cooperation with . . . Translation of *(or by)* . . . Presented at conference of . . . To be published in . . .

Block 19. Key Words. Select terms or short phrases that identify the principal subjects covered in the report, and are sufficiently specific and precise to be used as index entries for cataloging, conforming to standard terminology. The DoD "Thesaurus of Engineering and Scientific Terms" *(TEST)*, AD-672 000, can be helpful.

Block 20. Abstract. The abstract should be a brief *(not to exceed 200 words)* factual summary of the most significant information contained in the report. If possible, the abstract of a classified report should be unclassified and the abstract to an unclassified report should consist of publicly- releasable information. If the report contains a significant bibliography or literature survey, mention it here. For information on preparing abstracts see "Abstracting Scientific and Technical Reports of Defense-Sponsored RDT&E," AD-667 000.

★ U.S. GOVERNMENT PRINTING OFFICE : 1973—729-091/1431 3-1